Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

High order Ambisonic decoding method for irregular loudspeaker arrays

Jorge TREVINO (1,2), Takuma OKAMOTO (1,3), Yukio IWAYA (1,2) and Yôiti SUZUKI (1,2)

(1) Research Institute of Electrical Communication, Tohoku University, Sendai, Japan
(2) Graduate School of Information Sciences, Tohoku University, Sendai, Japan
(3) Graduate School of Engineering, Tohoku University, Sendai, Japan

PACS: 43.60.Sx, 43.60.Tj

ABSTRACT

Ambisonics, a sound field synthesis and reproduction technique, has shown promising results in conveying three-dimensional spatialized sound. Ambisonic encodings directly describe the spatial properties of sound fields without reference to the reproduction system. Precise regeneration of a sound field requires a large number of loudspeakers arranged so as to adequately sample all directions; this is referred to as a regular layout. It is not difficult to find a decoding matrix to reproduce Ambisonic recordings using uniformly distributed loudspeakers. Nevertheless, evenly locating a large number of loudspeakers is not feasible in most scenarios. Irregular arrays, however, are known to lead to ill-conditioned and singular re-encoding matrices. We propose a method for accurate reproduction of high order Ambisonic recordings over irregular loudspeaker arrays. Our approach consists of three stages which aim to exploit any regularities of the array, while using asymmetrically located loudspeakers to expand the listening volume. We evaluated our proposal using an irregular 157-channel loudspeaker array. Comparisons with mainstream decoding methods were conducted. The proposed scheme results in an overall increase in the size of the listening volume.

1. INTRODUCTION

Humans are sensitive to the spatial features of sound fields. We can accurately determine the direction from which sound reaches our ears, judge the distance to a given source, perceive the effects of obstacles in the sound propagation path, assess the properties of listening rooms through reverberation, and, to some extent, estimate the size and shape of sound sources [1]. A system intended to present a realistic auditory scene to its users must be able to convey all this information.

The most natural way to reproduce a three-dimensional sound space is to precisely recreate the original sound field. If the physical variables that humans perceive are reconstructed, a very realistic experience can be delivered. Unfortunately, sound fields carry too much information for exact reproduction to be practicable. A compromise must be made between the reproduction system's complexity and its accuracy. Since different applications demand varying degrees of precision, a scalable format for the characterization of sound fields is desirable.

Ambisonics is a promising technique that manages to encode a given sound field with arbitrary accuracy [2–4]. Ambisonic encodings focus on the physical properties of sound fields, making them independent of the reproduction system; they can be reproduced over a variety of loudspeaker arrays. We provide a brief introduction to the basics of Ambisonics and outline the most common Ambisonic decoding methods in Section 2. Standard decoders show excellent results when targeting arrays with regular loudspeaker layouts; however, the uniform distribution of many loudspeakers is often impracticable. Mainstream acceptance of Ambisonics calls for a decoder capable of using irregular arrays. Unfortunately, naïve approaches to the decoding of Ambisonic recordings for reproduction using irregular loudspeaker arrays can lead to numerical instability and suboptimal results. The difficulties of Ambisonic reproduction using irregular arrays are summarized in Section 3. Following the trajectory of our previous research [5–7], we present a method to decode high order Ambisonic streams. Our proposal, detailed in Section 4, integrates three decoding strategies for different loudspeaker distributions, so that irregularities in the array will not introduce significant numerical errors. A high order Ambisonic decoder for an irregular 157-channel loudspeaker array was developed using the proposed method. Section 5 compares its performance with that of a standard Ambisonic decoder.

2. AMBISONICS

Ambisonics relies on multiple audio channels to encode and reproduce a given sound field at the position of the listener. Sound fields can be either synthesized by simulating the propagation of sound from virtual sources or recorded using microphone arrays. The spatial features of sound fields are encoded as a set of expansion coefficients in an approach similar to how smooth functions can be characterized by their Taylor expansions or periodic functions by their Fourier series. Ambisonic recordings can be reproduced using loudspeaker arrays to reconstruct the encoded sound fields. An important advantage of the Ambisonic format is its independence from the particulars of the reproduction system. Any loudspeaker array with a proper decoder can reproduce Ambisonic recordings; however, the accuracy to which the sound field can be reconstructed depends on the loudspeaker distribution. We refer to arrays in which the loudspeaker positions regularly sample all directions as *regular loudspeaker arrays*. Configurations where all loudspeakers lie on the surface of a sphere centered at the listening position are designated as *spherical loudspeaker arrays*. Both properties are independent; an array can be regular or irregular irrespective of its being spherical or not.

2.1 Multipole expansion

Assuming an ideal propagation medium, a sound field p can be mathematically described by the scalar wave equation

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right] p(r,\theta,\varphi,t) = 0, \tag{1}$$

where θ and φ stand for the azimuthal and polar angles, respectively, r denotes the radial coordinate, t designates the temporal coordinate and c represents the speed of sound.

By expressing the time dependence in terms of normal modes, the general solutions to Eq. (1) can be written as [8]

$$p(\vec{r},k) = \left[\sum_{m=0}^{\infty} j_l(kr) \sum_{n=-m}^{m} B_{mn}(k) \mathrm{Y}_{mn}(\theta,\varphi)\right] e^{i(2\pi ck)t} \tag{2}$$

for all wavenumbers k, assuming the magnitude of $\vec{r}$ to be smaller than the distance to the closest sound source. The symbol j_l stands for the spherical Bessel function of order $l = m(m+1)$, and Y_{mn} denotes the spherical harmonic function of degree m and order n.

Eq. (2) is known as the multipole expansion of $p(\vec{r},k)$. The expansion coefficients

$$B_{mn}(k) = \int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{p(\vec{r},k)}{j_{m(m+1)}(kr)} \mathrm{Y}_{mn}(\theta,\varphi)\sin\varphi d\varphi d\theta, \tag{3}$$

constitute an Ambisonic encoding of the sound field.

Since sound fields are characterized by real-valued functions, it is useful to define the multipole expansion in terms of the real part of the spherical harmonic functions:

$$\mathrm{Y}_{mn}(\theta,\varphi) = \begin{cases} N_m^0 P_m^0(\cos\varphi) & \text{if } n = 0, \\ \sqrt{2} N_m^n P_m^n(\cos\varphi)\cos n\theta & \text{if } n > 0, \\ \sqrt{2} N_m^n P_m^{-n}(\cos\varphi)\sin n\theta & \text{if } n < 0, \end{cases} \tag{4}$$

where P_m^n represents an associated Legendre polynomial and N_m^n denotes a normalization constant.

The radial dependence of Eqs. (2) and (3) can be simplified if all the sound sources are located at distances $r \gg 1/k$. The asymptotic limit of the Bessel functions [9] yields what is known as the plane wave approximation. Eqs. (2) and (3) can be rewritten in this limit as

$$p^+(\hat{r},k) = \sum_{m=0}^{\infty}\sum_{n=-m}^{m} B_{mn}(k)\mathrm{Y}_{mn}(\theta,\varphi), \tag{5}$$

$$B_{mn}(k) = \int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} p^-(\hat{r},k)\mathrm{Y}_{mn}(\theta,\varphi)\sin\varphi d\varphi d\theta, \tag{6}$$

where $p^{\pm}(\hat{r},k)$ approximate the sound pressure field attenuated by $1/kr$ and phase shifted by $\pm 2\pi kr$.

2.2 Decoding of Ambisonic recordings

Ambisonic encodings of sound fields, given by Eqs. (3) and (6), are independent of the reproduction system. An Ambisonic decoder is required to translate them into signals for a specific loudspeaker array. Exact reconstruction of an arbitrary sound field using Eq. (2) or Eq. (5) requires an infinite number of expansion coefficients. In practice, the sum is truncated up to a given degree, known as the Ambisonic order. Naturally, by increasing the Ambisonic order, greater reconstruction accuracy is attained.

Basic decoding of Ambisonic data for reproduction using a loudspeaker array is achieved by computing a weighed sum of all Ambisonic channels for each loudspeaker. Weights can be calculated by first defining a re-encoding matrix

$$\mathbf{B}(k) = \mathbf{C}\mathbf{p}(k). \tag{7}$$

The components of vector $\mathbf{p}(k)$ are the loudspeaker signals. The re-encoding matrix $\mathbf{C}$, therefore, has $(N+1)^2$ rows and one column for every loudspeaker in the array. The elements of $\mathbf{C}$ are given by the spherical harmonic functions evaluated at the directions of the loudspeakers, (θ_s,φ_s), as follows:

$$c_{(m+1)^2-m+n,s} = \mathrm{Y}_{mn}(\theta_s,\varphi_s). \tag{8}$$

By comparison with Eq. (6), the re-encoding equation, Eq. (7), can be interpreted as the Ambisonic encoding of the sound field generated by the array when the loudspeakers are fed signals $\mathbf{p}(k)$.

Loudspeaker signals corresponding to the Ambisonic encoded sound field can be computed by solving the linear system of Eq. (7). For the re-encoding matrix to be invertible, however, the number of loudspeakers in the array must match the count of Ambisonic channels to be decoded. In practice, it is desirable to use larger arrays to improve the reproduction accuracy. It is common to rely on the Moore-Penrose pseudoinverse when designing an Ambisonic decoder. The decoding equation can be written in terms of the pseudoinverse of $\mathbf{C}$, denoted by $\mathbf{C}^+$, as [10]

$$\mathbf{p}(k) = \mathbf{C}^+\mathbf{B}(k). \tag{9}$$

If the number of loudspeakers in the array is larger than the quantity of Ambisonic channels, Eq. (9) minimizes the Euclidean norm of $\mathbf{p}(k)$. If the array has fewer loudspeakers than the number of Ambisonic channels, accurate reconstruction becomes impossible in general; however, Eq. (9) will result in the loudspeaker signals minimizing the Euclidean norm of the error vector $\mathbf{C}\mathbf{p}(k) - \mathbf{B}(k)$ [11].

2.3 Loudspeaker distance compensation and near field effects

Reproduction of Ambisonic recordings using Eq. (9) does not take into consideration the finite distance between the listener and the loudspeakers. When using non-spherical arrays, that is, arrays where the loudspeakers are not equidistant from the listening position, additional processing of the loudspeaker signals is required.

By expressing the complete system's amplification and propagation time in terms of a reference distance r_{ref}, Eqs. (5) and (6) show that the $s-$th loudspeaker requires its signal to be amplified by a factor of r_s/r_{ref} and phase shifted by $(r_{\mathrm{ref}} - r_s)\omega/c$. Causality is ensured by making r_{ref} greater than or equal to the distance between the listening position and the farthest loudspeaker in the array.

The plane wave approximation only holds for arrays where the distance between the listener and the loudspeakers is much larger than the wavelength of the sound signal. Particularly at low frequencies, this requires very vast arrays. Furthermore, the spherical Bessel functions approach their asymptotic limit more slowly as the order is increased, compromising the accuracy of high order Ambisonic systems. Relatively small arrays require a compensation stage to cancel the near field effects, that is, those effects that are not captured by the plane wave approximation.

Near field effects are contained in Eqs. (2) and (3). Radial dependence can be introduced into the re-encoding matrix of Eq. (7) as follows:

$$\mathbf{B}(k) = \mathbf{F}(kr_s)C\mathbf{p}(k). \tag{10}$$

Here, $\mathbf{F}(kr_s)$ is a diagonal, square matrix whose dimensions are equal to the number of Ambisonic channels, and its elements are given by

$$f_{ij}(kr_s) = j_{i(i+1)}(kr_s)\delta_{ij}, \tag{11}$$

where δ_{ij} denotes the Kronecker delta.

Near field effects can be suppressed by filtering the Ambisonic stream with the filters $\mathbf{F}^{-1}(kr_s)$ before the standard decoding of Eq. (9). These filters are always stable [12], but depend explicitly on the distance between the loudspeakers and the listening position. Therefore, non-spherical loudspeaker arrays must process the Ambisonic recordings using a filter bank with filters for every distinct distance between the listener and a loudspeaker.

3. DIFFICULTIES WHEN DECODING FOR IRREGULAR LOUDSPEAKER ARRAYS

The decoding equation, Eq. (9), is completely general and does not impose any restrictions on the geometry of the loudspeaker array. However, by relying on the pseudoinverse, its practical applicability is limited. There are two main difficulties with this approach, particularly when decoding for an irregular loudspeaker array: numerical instability and suboptimal solutions. It is very difficult to build arrays with loudspeakers evenly distributed in every direction. By considering the full range of spherical harmonic functions up to a given degree, the decoding equation might force the target system to try to reconstruct a sound field it cannot physically recreate.

3.1 Mixed-order Ambisonics

When building a loudspeaker array, distributing a large number of loudspeakers horizontally is usually easier than positioning them at varying elevations. Humans also show greater spatial accuracy when localizing sound sources in the horizontal plane [1]. It is, thus, useful to characterize sound fields with more emphasis on their horizontal features.

An initial approach to improve the reproduction of Ambisonic recordings using arrays with irregular configurations is to ignore the expansion coefficients corresponding to spherical harmonic functions that are not adequately sampled by the positions of the loudspeakers. Decoding for arrays with different horizontal and vertical spatial resolutions can be achieved by considering the multipole expansion up to a given degree, and complementing it with the expansion coefficients corresponding to the horizontally oriented spherical harmonics of higher degrees. This scheme is known as mixed-order Ambisonics.

There are exactly two horizontally oriented spherical harmonic functions for every non-zero degree: those for which the absolute value of the order is equal to the degree. The re-encoding matrix given by Eq. (8) can be rewritten for a mixed-order Ambisonic system as

$$\begin{aligned} c_{(m+1)^2-m+n,s} &= Y_{mn}(\theta_s,\varphi_s) && \text{for } m < N_{FS}, \\ \left.\begin{aligned} c_{N_{FS}^2+2m,s} &= Y_{m,-m}(\theta_s,\varphi_s) \\ c_{N_{FS}^2+2m+1,s} &= Y_{mm}(\theta_s,\varphi_s) \end{aligned}\right\} & && \text{for } m > N_{FS}, \end{aligned} \tag{12}$$

where N_{FS} denotes the Ambisonic full-sphere order, that is, the degree up to which all of the multipole expansion terms are considered.

Decoding for mixed-order Ambisonic systems can be done by discarding the components of $\mathbf{B}(k)$ that are not represented in the re-encoding matrix of Eq. (12) and applying the techniques discussed in the previous section.

3.2 Numerical instability of the decoding equation

Mixed-order Ambisonics facilitates the use of arrays with distinct horizontal and vertical loudspeaker distributions. The mixed-order decoding process, however, still relies on the pseudoinverse of a re-encoding matrix.

While the pseudoinverse provides least squares solutions to underdetermined and overdetermined linear systems, it is not a continuous operation. Slight changes to a matrix can drastically alter its pseudoinverse. Such behavior can be understood from the singular value decomposition [11]. A matrix with a singular value equal to zero will be very slightly changed by turning it into a small positive number; however, the pseudoinverse will be modified by the reciprocal of this small value, resulting in a very large change. When doing infinite precision arithmetic, such problems can be generally ignored. Nevertheless, practical implementation demands finite precision. A first approach to overcome this difficulty is to discard values smaller than a properly chosen threshold, dependent on the numerical precision of the computation system, when evaluating the pseudoinverse.

Replacing small singular values with zeros before evaluating the pseudoinverse prevents the strong amplification of round-off errors. There are, however, other sources of error in the Ambisonic reproduction process. It is impossible to exactly measure the layout of a physical array. Moreover, Ambisonic recordings will invariably contain some level of noise, audio systems such as DACs and amplifiers do not possess ideal frequency responses, loudspeakers are not perfect omnidirectional radiators, the most fundamental premise of the Ambisonic strategy, Eq. (1), neglects non-linear effects, dissipation and other properties of a real medium, and perturbations due to the presence of the listener inside the array, as well as room reverberation, are not considered.

Since it is impossible to eliminate every source of error, a way to evaluate the feasibility of decoding through Eq. (9) is needed. The numerical stability of solutions relying on the pseudoinverse of a matrix $\mathbf{C}$ can be judged from its condition number, defined as the norm of the ratio between the maximal and minimal singular values of $\mathbf{C}$. Alternatively, it can be calculated as

$$\operatorname{cond}(\mathbf{C}) = \|\mathbf{C}\|\|\mathbf{C}^+\|, \tag{13}$$

where $\|\cdot\|$ denotes a matrix norm. A large condition number implies that $\mathbf{C}$ is ill-conditioned and least squares solutions are numerically unstable.

In general, re-encoding matrices derived for regular loudspeaker arrays possess smaller condition numbers [13]. This reflects the fact that regular configurations can more accurately sample the spherical harmonic functions. Therefore, accurate decoding of Ambisonic data using Eq. (9) requires a loudspeaker array with a sufficiently regular layout.

3.3 Suboptimal solutions

Decoding matrices derived using the pseudoinverse can result in suboptimal reproduction of Ambisonic recordings. It is common for loudspeaker arrays to possess more loudspeakers than the number of available Ambisonic channels; thus, their re-encoding matrices correspond to underdetermined linear systems. The least squares solutions provided by the pseudoinverse for underdetermined systems are those with minimal Euclidean norm [11]. Minimizing the Euclidean norm of the loudspeaker signals does not ensure the best results from the perspective of a human listener.

The multipole expansion given by Eqs. (2) and (3) characterizes

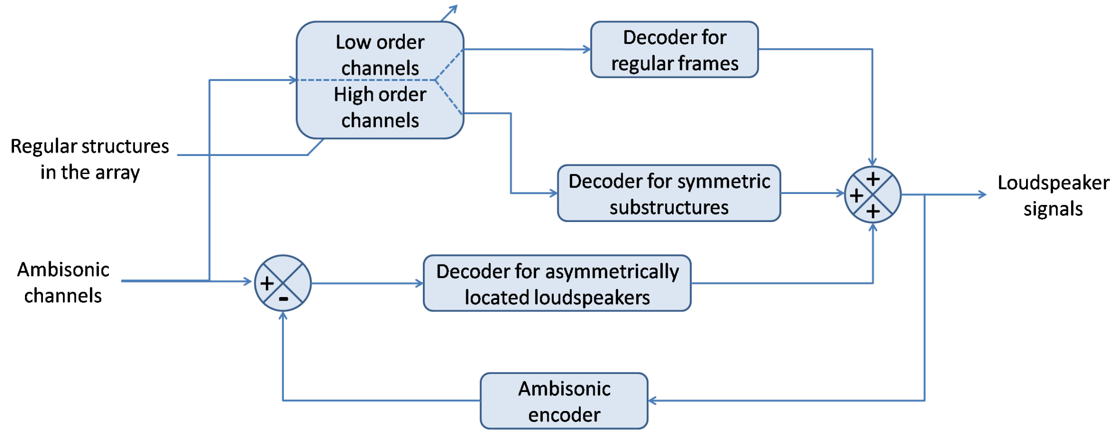


Figure 1: Block diagram of the proposed Ambisonic decoder for irregular arrays

the spatial dependencies of sound fields as seen from the origin. The hearing experience of a human listener, however, is not bound to a single spatial point. The listening volume, that is, the region within the array where the reconstruction error remains low, must be large enough to accommodate a human listener. Furthermore, to ensure a comfortable experience, the listener's movement should not be restricted to a very small spatial region. Perfect reconstruction of an Ambisonic-encoded sound field at the origin is, therefore, not enough to guarantee the best results.

Loudspeaker signals calculated using Eq. (9) minimize the reconstruction error at the center of the array; however, there are no additional constraints on the spatial distribution of the error. The approximate rotational invariance of loudspeaker arrays with uniform layouts helps to maintain a low reconstruction error around the origin, even if the decoding stage is based on the pseudoinverse of a re-encoding matrix. The loudspeaker distribution of an irregular array, however, does not consistently sample all of the spherical harmonic functions. Reconstruction accuracy over an extended region when using an irregular array cannot be guaranteed by minimizing the error at a single point.

4. A NEW AMBISONIC DECODING METHOD SUITABLE FOR IRREGULAR LOUDSPEAKER ARRAYS

Decoding of Ambisonic data through the pseudoinverse of a re-encoding matrix can lead to the drastic amplification of errors when targeting an irregular loudspeaker array. Another disadvantage of this method lies in its focus on a single spatial point: the center of the array. Mixed-order Ambisonics ameliorates the effects of distinct horizontal and vertical arrangements of loudspeakers by eliminating high order non-horizontal Ambisonic channels. Similar strategies can be devised for irregularities along other directions.

There is no fundamental reason to favor a decoder which simultaneously handles all Ambisonic channels or which uses all of the available loudspeakers to recreate every degree and order of the multipole expansion. Indeed, compensation for near-field effects already demands a collection of separate decoding stages when targeting non-spherical loudspeaker arrays. We propose that by separately decoding subsets of Ambisonic channels using adequate partitions of the loudspeaker array the limitations of the pseudoinverse can be surmounted.

Our proposal is summarized in the block diagram of Fig. 1. Ambisonic decoding is performed in three stages. In the first stage, an attempt is made to exploit any regular structure available within the layout of the loudspeaker array. If successful, the first stage allows decoding of the lower Ambisonic orders. Components that could not be decoded in the first stage are fed to a second decoder based on the mixed-order Ambisonics approach. The underlying discrete symmetries of the array are employed to decode as many terms of the multipole expansion as possible while preserving numerical stability. Finally, the asymmetric portions of the array are used to stabilize the radial dependence of the reconstruction error, increasing the listening volume.

Ambisonic recordings preserve the invariance under rotations and boosts of Eq. (1). Additional discrete symmetries are present in the Ambisonic data since the multipole expansion is limited to a finite degree. Irregular loudspeaker arrays, by definition, cannot be expected to possess such symmetries. The decoder must decompose the Ambisonic stream in accordance to the layout of the target loudspeaker array.

4.1 Decoding for regular frames

If the target array contains a set of loudspeakers arranged as the faces of a Platonic solid, accurate reconstruction of low Ambisonic orders can be achieved through Eq. (9). Specifically, a tetrahedral, hexahedral or octahedral substructure allows for the decoding of the first Ambisonic order using the full array. A dodecahedral substructure permits the decoding to be extended up to the second Ambisonic order. An icosahedral substructure can be used to decode Ambisonic data up to the third Ambisonic order.

4.2 Decoding for symmetric subsets of loudspeakers

The Ambisonic channels decoded thanks to a regular frame within the array, if available, should be discarded. The target array can now be scanned for discrete symmetries similar to those possessed by the spherical harmonic functions of degrees equal to the remaining Ambisonic orders. The symmetry groups of high degree spherical harmonic functions contain as subgroups the transformations that leave lower degree spherical harmonics invariant. Therefore, it is convenient to scan the array starting with the symmetries corresponding to the highest Ambisonic order available.

Once a useful, symmetric subset of loudspeakers has been found, a mixed-order Ambisonic re-encoding matrix is generated for it. The re-encoding matrix should cover all of the degrees and orders of the multipole expansion that are adequately represented by the frame. Decoding can be done using the procedures outlined in previous sections.

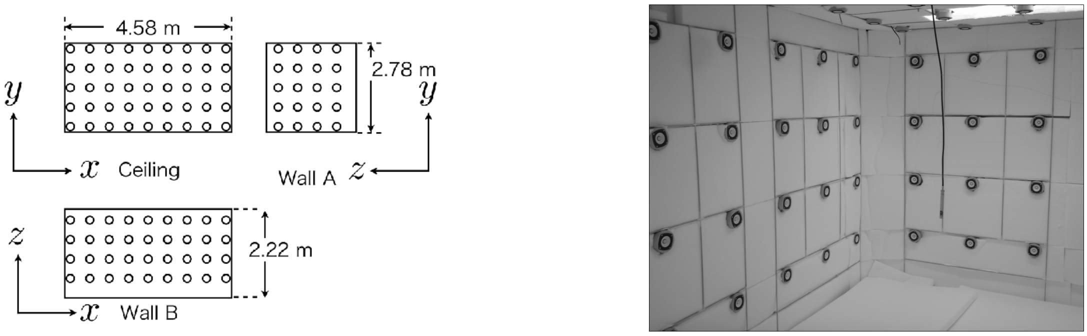


Figure 2: (a) Layout of the 157-channel irregular loudspeaker array. (b) 157-channel irregular loudspeaker array inside an anechoic room.

4.3 Decoding for asymmetrically located loudspeakers

Thus far, the proposed decoding method is just a generalization of the mixed-order Ambisonics approach. By decoding only those channels that correspond to properly-sampled spherical harmonic functions, numerical instability can be avoided. However, irregular arrays are seldom composed exclusively of regular substructures. Decoding for asymmetrically located loudspeakers cannot rely on the pseudoinverse.

Instead of seeking least squares solutions that minimize the Euclidean norm of the loudspeaker signals, a useful constraint on the solutions is to minimize the radial derivative of the reconstruction error

$$\left| \frac{\partial}{\partial r} \left[\tilde{\psi}(r,\theta,\varphi) - \phi(r,\theta,\varphi) - \sum_s \sum_{m=0}^{N} \sum_{n=-m}^{m} G^s_{mn}(k) \frac{e^{-ik|r-r_s|}}{|r-r_s|} \mathrm{Y}_{mn}(\theta_s,\varphi_s) \right] \right|, \tag{14}$$

where $\tilde{\psi}(r,\theta,\varphi)$ represents the sound field encoded by the Ambisonic stream, $\phi(r,\theta,\varphi)$ stands for the sound field reconstructed by the previous two stages. The first sum runs over all of the loudspeakers that are not part of a regular frame or symmetric subset. Gains $G^s_{mn}(k)$ minimizing Eq. (14) can be used to decode the Ambisonic data for asymmetrically located loudspeakers through the following decoding equation:

$$p_s(k) = [\mathbf{G}^s(k)]^{\mathrm{T}} \mathbf{B}(k). \tag{15}$$

Near field corrections can be made with filters defined by the reciprocal of Eq. (11). By minimizing the radial derivative of the reconstruction error, the proposed scheme increases the listening volume.

5. IMPLEMENTATION

Two fifth order Ambisonic decoders for an irregular, 157-channel loudspeaker array were developed. The first one is a conventional decoder based on Eq. (9), while our proposed decoding method was employed for the second one. Both decoders included a stage to correct for near-field effects. Ambisonic encodings of plane waves were synthesized and used to compare the performance of both decoders.

5.1 157-channel loudspeaker array

The target array used for demonstration purposes is the one described in [5]. It consists of 157 loudspeakers arranged on the walls and ceiling of a rectangular room. The array is 4.58 m deep, 2.78 m wide and 1.72 m high. The listener is assumed to be located at the center, facing in the direction of the positive $x-$axis. The distribution of the loudspeakers is displayed in Fig. 2a. The walls, floor and ceiling of the room housing the array are covered by 3 layers of 5 cm thick sound absorbing material made from a non-woven fabric [14]. The room presents an almost anechoic response above 125 Hz. A photograph of the array is shown in Fig. 2b. Conventional Ambisonic decoding targeting this particular array is numerically unstable for orders higher than five [7].

5.2 Proposed decoder for the irregular 157-channel loudspeaker array

We developed an Ambisonic decoder for the 157-channel array by applying the strategy described in Section 4. The decoding for symmetric subsets of loudspeakers was restricted to the cyclic groups defined over the horizontal plane and the two planes of symmetry of the target array. This might not cover all of the transformations leaving both, the arrangement of the loudspeakers and the spherical harmonic functions invariant. Further research into the symmetries of high order Ambisonic streams might lead to performance improvements in this stage.

Decoding gains for asymmetrically located loudspeakers were approximated by first calculating the residual sound field $[\tilde{\psi}(r,\theta,\varphi) - \phi(r,\theta,\varphi)]\, e^{i(2\pi ck)t}$ through a computer simulation. Specifically, $\phi(r,\theta,\varphi)$ was derived by assuming the loudspeakers to be ideal monopole radiators. The reconstructed sound field was evaluated at 36 points lying over a sphere located in the center of the array. The radius of the sphere was chosen in terms of a sampling frequency of 48 kHz, as $r_{\mathrm{sphere}} = c/f_s \approx 7.08$ mm. The sampling points were distributed in a Fliege geometry [15]. An Ambisonic encoding of the reconstructed sound field was calculated using Eq. (3). The residual sound field was then obtained by subtracting the original and reconstructed Ambisonic encodings and using them to evaluate Eq. (2). The residual sound field was time reversed and propagated from the sampling sphere to the locations of the loudspeakers. A second time reversal was performed to derive new loudspeaker signals which compensate for the residual sound field. While this approach does not ensure gains $G^s_{mn}(k)$ minimizing Eq. (14), it offers a good approximation if the sampling frequency used in the simulation is assumed to be much larger than the frequencies present in the sound field.

5.2 Results

The reconstruction accuracy of both decoders was calculated over a region spanning 30% of the total volume of the loudspeaker array. All results were obtained through a computer simulation, assuming the loudspeakers to be ideal omnidirectional radiators. Four different sound fields, consisting of single plane waves, were reconstructed. Two frequencies (500 Hz and 1 kHz) and two incident angles ($\theta = 0$ and $\theta = 30°$) were considered. Simulation results are shown in Figs. 3 to 10. Overall, the new decoding method outperforms the pseudoinverse method from the listener's perspective. While the proposed decoder is less accurate at the precise center of the array, the reconstruction error remains low throughout a larger volume.

Figs. 3a to 3c show the simulation results when reconstructing a 500 Hz plane wave incident from the right (positive x axis). Both decoders perform acceptably in this scenario; however, the wavefront produced by the new decoding method has a smaller curvature. Reconstruction error is shown in Figs. 4a and 4b. Conventional decoding results in very high reconstruction accuracy at the precise center of the array; nevertheless, slight displacements result in significant error variations. Our proposal shows reduced accuracy at the center, but the error is not significantly large. On the other hand, the new decoding method maintains low error levels throughout a larger volume.

Figs. 5a to 5c show the Ambisonic reconstruction of a 500 Hz plane wave incident at an azimuth angle of 30 degrees. The uneven distribution of the loudspeakers results in less accurate reconstruction than that achieved for the right-incident plane wave. Nevertheless, reconstruction error remains at an acceptably low level. The difference between the conventional, pseudoinverse-based decoder and the proposed decoding method is very significant. Reconstruction errors are shown in Figs. 6a and 6b. Different scales are used in each figure in order to convey the structure of the error distributions; using equal scales, with light and dark regions denoting large and small errors respectively, would result in an almost completely black image for the error achieved by our proposal, or a blank figure for the reconstruction error when using the pseudoinverse method. Not only is reconstruction accuracy significantly improved by our decoding scheme, a considerably large listening volume is also observed.

Figs. 7a to 7c show the results for a right-incident, 1 kHz plane wave. Ambisonic reconstruction of high frequency sources requires greater spatial resolution. Since the number of loudspeakers in the array is fixed, the reconstruction error, shown in Figs. 8a and 8b, is greater than that of the 500 Hz examples. The performance of the pseudoinverse decoder suffers considerably from the increase in frequency, resulting in a very limited listening volume. Our proposal also shows significant performance degradation; however, the listening volume can be seen to be significantly larger than that attained by the pseudoinverse decoder.

Finally, Figs. 9a to 9c show the reconstruction of a 1 kHz plane wave incident at an azimuth angle of 30 degrees. The accuracy of both decoders is inferior to that achieved in the other three scenarios. The reasons for the performance degradation are the higher spatial resolution required to characterize higher frequencies and the uneven layout of the loudspeaker array. The reconstructed wavefronts show curvatures similar to those observed for the 1 kHz right-incident plane wave. The differences between the original and reconstructed sound fields are shown in Figs. 10a and 10b. Reconstruction based on the pseudoinverse results in a considerably small listening volume, even as the reconstruction error at the center remains low. Our method results in a more extended region over which the error does not vary significantly.

6. CONCLUSIONS

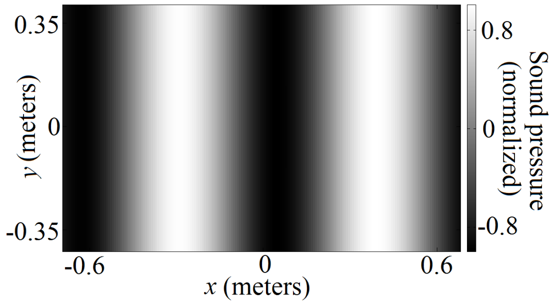
A new Ambisonic decoding method for irregular loudspeaker arrays was proposed and evaluated. Starting from the wave equation, and specifying the usual decoding techniques, two important drawbacks to the reliance on the pseudoinverse for decoding were identified. An alternate method, free of these drawbacks, was advanced. The proposal was evaluated using an irregular 157-channel loudspeaker array. The proposed method outperforms the conventional decoder from the perspective of the listener, due to an increased listening volume.

ACKNOWLEDGEMENTS

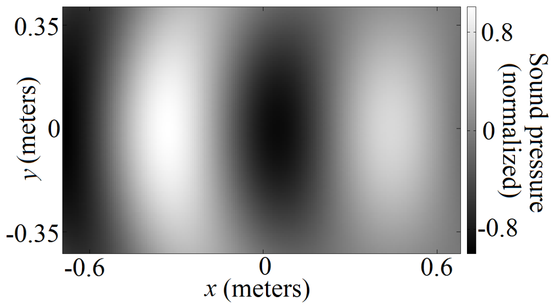
This study was partly supported by the GCOE program (CERIES) of the Graduate School of Engineering, Tohoku University and Grant-in-Aid of JSPS for Specially Promoted Research (no. 19001004).
The anechoic condition of the room housing the loudspeaker array was improved and measured with the valuable help of Dr. Densil Cabrera, member of the Faculty of Architecture, Design and Planning at the University of Sydney.

REFERENCES

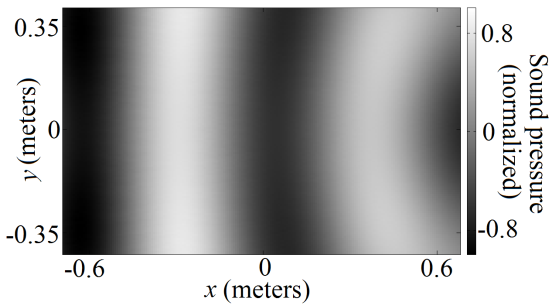
[1] J. Blauert, *Spatial Hearing*, MIT Press, Revised edition, 1996.
[2] M.A. Gerzon, "Periphony: With-Height Sound Reproduction," *J. Audio Eng. Soc.*, Vol. 21, no 1, pp. 2–10, Feb. 1973.
[3] D. Malham, "Higher order Ambisonic systems. Space in Music - Music in Space," *M.Phil. Thesis, University of York*, Apr. 2003.
[4] M.A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics, " *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, Nov. 2005.
[5] T. Okamoto, B.F.G. Katz, M. Noisternig, Y. Iwaya, Y. Suzuki, "Implementation of real-time room auralization using a surrounding loudspeaker array," *International Workshop on the Principles and Applications of Spatial Hearing (eProceedings)*, No. 10.1142-9789814299312-0049, Nov. 2009.
[6] J. Trevino, T. Okamoto, Y. Iwaya, Y. Suzuki, "Development of a higher order ambisonics encoder and decoder for a 157-channel surround speaker array," *Proc. 5th International Symposium on Medical, Bio- and Nano-Electronics*, pp. 137–138, Feb. 2010.
[7] T. Okamoto, D. Cabrera, M. Noisternig, B.F.G. Katz, Y. Iwaya, Y. Suzuki, "Improving Sound Field Reproduction in a Small Room Based on Higher-Order Ambisonics with a 157-Loudspeaker Array," *Proc. 2nd International Symposium on Ambisonics and Spherical Acoustics*, May 2010.
[8] J.D. Jackson, *Classical Electrodynamics, Third Edition*, ed. Wiley, 1998.
[9] E. Merzbacher, *Quantum Mechanics, Third Edition*, ed. Wiley, 1997.
[10] J. Daniel, "Acoustic field representation, application to the transmission and the reproduction of complex sound environments in a multimedia context," *Ph.D. Thesis, Université Paris*, 2000.
[11] G.H. Golub, C.F. Van Loan, *Matrix Computations, Third Edition*, ed. Baltimore, 1996.
[12] J. Daniel, "Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format," *Proc. AES 23rd International Conference*, May 2003.
[13] F. Zotter, "Sampling Strategies for Acoustic Holography/Holophony on the Sphere," *Tech. Rep. Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz*, 2008.
[14] Y. Suzuki, F. Saito, K. Nakai, H. Shimada, "Application of Polyester Nonwoven Fabric as Absorbing Material for Wedges in an Anechoic Room," *Proc. internoise*, No.382, pp. 1-6, 2004.
[15] J. Fliege, U. Maier, "The distribution of points on the sphere and corresponding curvature formulae," *IMA J. Numer. Anal.*, vol. 19, pp. 317-334, 1999.



(a) Ideal plane wave.

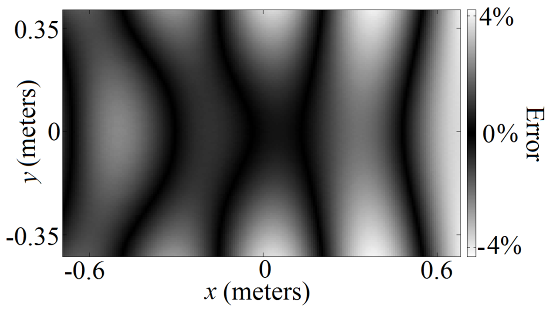


(b) Reconstruction using the pseudoinverse method.

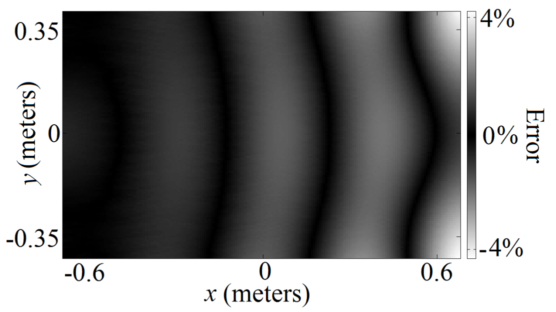


(c) Reconstruction using the proposed method.

Figure 3: 5th order Ambisonic reconstruction of a 500 Hz plane wave incident from the right.

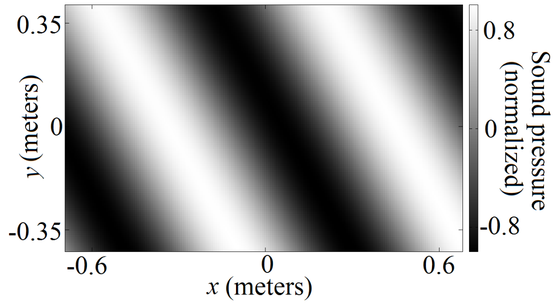


(a) Reconstruction error when using the pseudoinverse method.

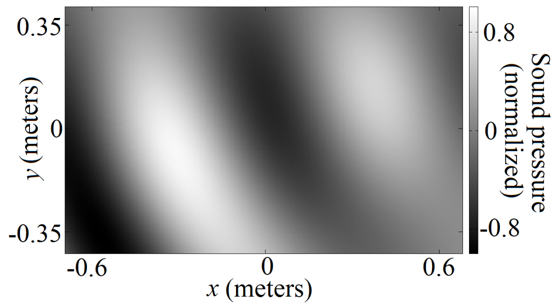


(b) Reconstruction error when using the proposed method.

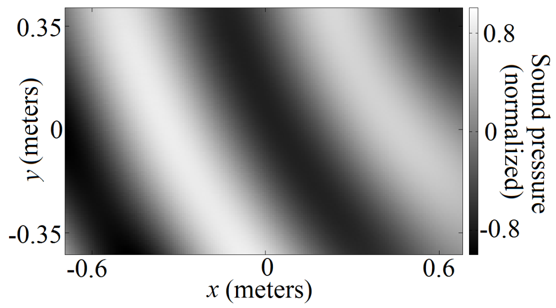
Figure 4: Reconstruction error for a 500 Hz plane wave incident from the right.



(a) Ideal plane wave.

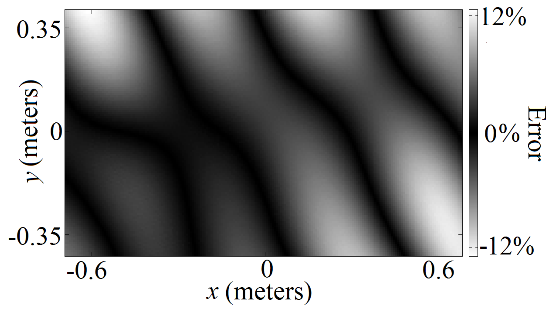


(b) Reconstruction using the pseudoinverse method.

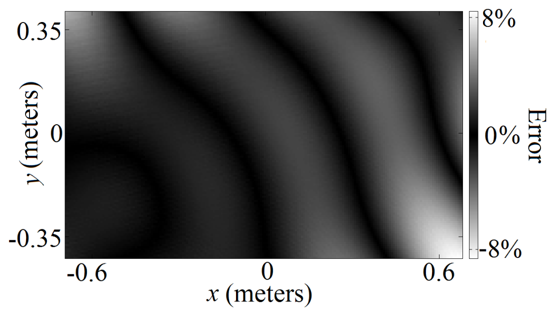


(c) Reconstruction using the proposed method.

Figure 5: 5th order Ambisonic reconstruction of a 500 Hz plane wave incident at an azimuth of 30°.

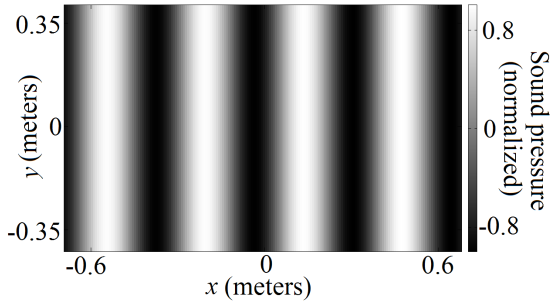


(a) Reconstruction error when using the pseudoinverse method.

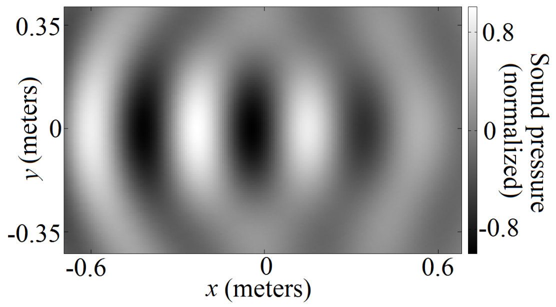


(b) Reconstruction error when using the proposed method.

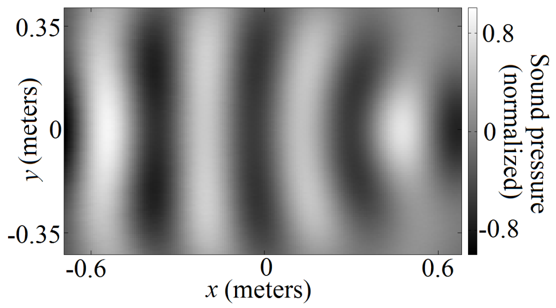
Figure 6: Reconstruction error for a 500 Hz plane wave incident at an azimuth of 30°.



(a) Ideal plane wave.

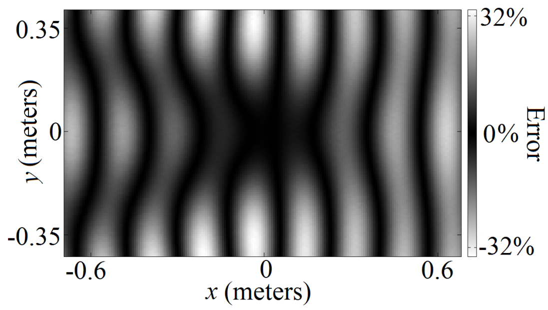


(b) Reconstruction using the pseudoinverse method.

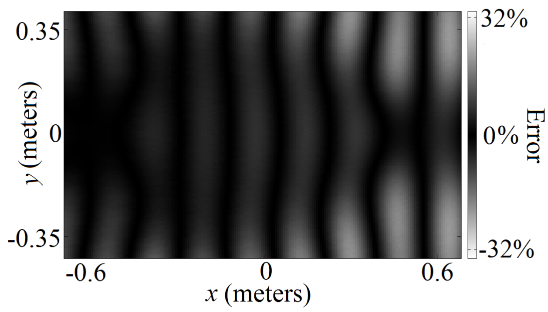


(c) Reconstruction using the proposed method.

Figure 7: 5th order Ambisonic reconstruction of a 1 kHz plane wave incident from the right.

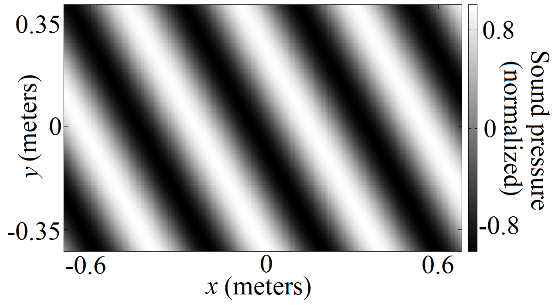


(a) Reconstruction error when using the pseudoinverse method.

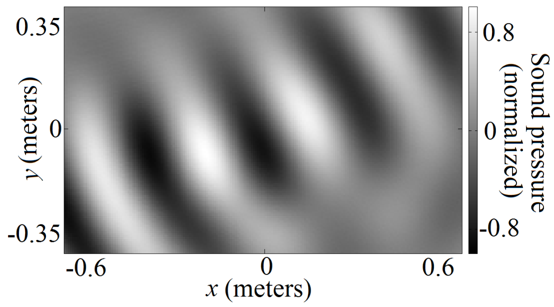


(b) Reconstruction error when using the proposed method.

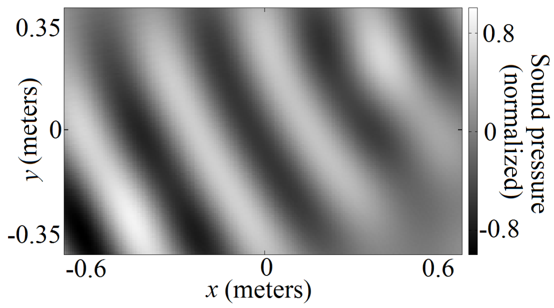
Figure 8: Reconstruction error for a 1 kHz plane wave incident from the right.



(a) Ideal plane wave.

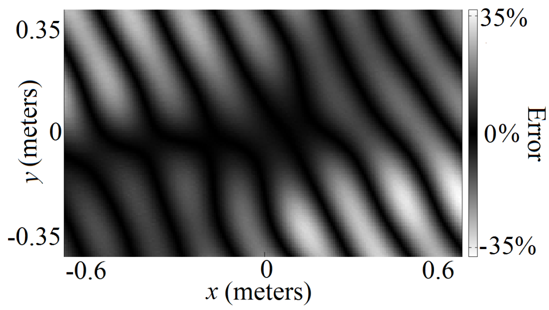


(b) Reconstruction using the pseudoinverse method.

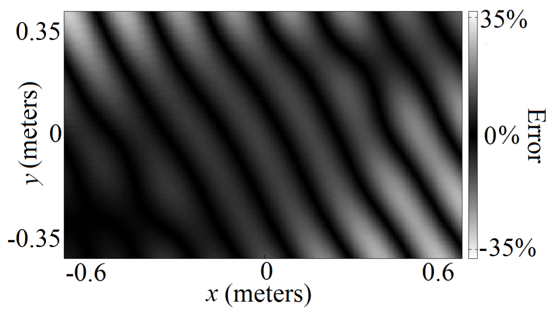


(c) Reconstruction using the proposed method.

Figure 9: 5th order Ambisonic reconstruction of a 1 kHz plane wave incident at an azimuth of 30°.



(a) Reconstruction error when using the pseudoinverse method.



(b) Reconstruction error when using the proposed method.

Figure 10: Reconstruction error for a 1 kHz plane wave incident at an azimuth of 30°.